

データサイエンス教育支援のための リアルタイムオブジェクト検出モデルの設計ツール

永田研究室 F122021 黒住 旺生

1. 目的

デジタル化が加速する現代社会ではAIを利用するだけでなく、AIがどのようなデータから構築され、どのようなプロセスで予測や自動生成を可能にしているのかを理解しておくことも求められている。本学では日常生活、社会生活の現場でAIなどを使いこなすための数理・データサイエンスの基礎的素養を身につけるための教育プログラムが提供されており、実験的に体験学習できる教育システムの構築も重要となっている。本研究では自動運転、品質管理、医療画像診断、移動ロボットなどへの応用が進んでいるリアルタイム物体検出モデル YOLO のための学習支援システムの構築を目指している。

2. 研究内容

本研究で取り扱うリアルタイムオブジェクト検出とは、深層学習を用いて動画やカメラ映像内の複数の物体を毎秒数十フレーム以上の処理速度で、クラス名と位置をそれぞれリアルタイムに識別、特定する。実験ではdesk, chair, door, person, alcohol, wheel_chair, wheel_person, vending_machineの8クラスのオブジェクトが含まれた画像を撮影し、350枚程度収集した。リアルタイム性を考慮し、画像サイズを480×853に縮小させ、MATLAB上で利用できるアノテーションソフト「イメージラベラー」を用いて矩形のバウンディングボックススペースで8クラスのラベリング処理を行った。画像を増やしながら追加でラベリングを行い、gTruthファイルに保存し、tiny-cocoをバックボーンとしたYOLOXモデルの学習に用いた。研究室で開発したツールを用いて保存したgTruthファイルの選択から、エポック数、ミニバッチ数、学習率などの設定までを行う。エポック数は画像枚数に応じて値を変更する。訓練後のYOLOXモデルの汎化性能は、テスト用の画像や動画をリアルタイムで物体検出することで、検出されたクラスの数、クラス検出時のスコアから評価を行うことができる。少ない数のラベリングオブジェクトから徐々に数を増やすことで汎化性能の向上を体験できるのかを検証する。

3. 結果

訓練に用いたラベル付き画像の増加に伴う各クラスの検出スコアの平均値(Mean score)と誤検出(Over detection)された画像枚数を表1に示す。ここでの誤検出とは、リアルタイムオブジェクト検出を行った際に動画内には存在していなかったクラスが誤って検出されてしまうことである。表1より訓練に用いたラベル画像の数が異なる9種類のYOLOXモデルで汎化性能を比較したところ、各クラスの検出スコアの変化にばらつきがあるものの、全体的な汎化性能の向上が確認された。提案システムを用いることで、体験学習に参加した学生は汎化性を高めながらYOLOXモデルを構築していくために必要となる、ラベリング処理で作成したgTruthファイルを使った訓練からオブジェクト検出の評価といった一連の流れを効率的に理解できるものと期待される。今後は、移動ロボットなどのハードウェアに実装することで、フィジカルAIとしてのリアルタイムオブジェクト検出に取り組んでいきたい。

Table 1 Evaluation of nine kinds of YOLOX models on real-time object detection.

Number of images	21	40	80	121	169	223	261	301	350
Label									
person	0.74	0.82	0.78	0.79	0.82	0.84	0.81	0.88	0.90
desk	0.74	0.77	0.69	0.90	0.86	0.83	0.76	0.80	0.87
chair	0.81	0.64	0.81	0.82	0.90	0.76	0.81	0.82	0.84
door	0.61	0.59	0.54	0.66	0.65	0.70	0.68	0.71	0.77
alcohol	0.00	0.61	0.61	0.59	0.64	0.80	0.82	0.84	0.81
wheel_chair	0.78	0.82	0.75	0.79	0.74	0.79	0.87	0.88	0.87
person_wheel	0.66	0.63	0.73	0.72	0.63	0.71	0.74	0.71	0.72
vending_machine	0.79	0.79	0.89	0.77	0.86	0.86	0.86	0.90	0.91
Mean score	0.64	0.71	0.72	0.76	0.76	0.79	0.79	0.82	0.84
Over detection	48	27	74	84	42	6	33	14	16